REACTION-TIMES AS A TEST OF MENTAL ABILITY.[1]

By GUY MONTROSE WHIPPLE,

Assistant Professor of the Science and Art of Education,
Cornell University.

This paper is an attempt to show (1) that a distinction must be made between what may be called the 'laboratory' and the 'anthropometric' types of reaction experiment, (2) that many of the reaction-tests conducted upon school-children have not conformed to the usual requirements of the 'laboratory' method, and (3) that reaction-time tests, of whatever type, cannot be successfully used in tests of school-children, and fail to indicate mental ability. I shall not try in what follows to keep these propositions sharply distinct, or to keep to the order just given.

The interest in reaction-times as studied in the earlier investigations (incited by the phenomena of the 'personal equation' and Wundt's complication experiments during the early sixties) was directed very largely to the determination of the time-values themselves. But it was soon discovered that the results of different investigators failed to evince conspicuous harmony, even when the outward conditions seemed entirely identical. Thus attention was turned to the qualitative analysis of the psychical or psychophysical conditions governing the process of reaction. The treatment became qualitative rather than quantitative. Naturally many factors were unearthed,—some of them but little suspected in the previous work. It was found, for instance, that the simple reaction-time depended not only upon such objective factors as the modality of the stimulus, its extent and intensity, upon the nature of the movement of reaction, etc., but also quite as much upon certain more subjective conditions, such as the degree of fore-knowledge or expectant attention aroused in the observer, the presence or absence of a warning-signal, the length and uniformity of the interval between this signal and the stimulus, the presence or absence of distraction, the general condition of the nervous system as affected by drugs, fatigue products, etc., and, perhaps most significant of all, upon the direction of the observer's attention,—an influence cited by Orschansky in 1887, but still better known by the work of Lange soon after published.[2]

[1] A paper read at the Meeting of Experimental Psychologists at Ithaca, in April, 1904.

[2] *Philos. Studien*, IV, 1888, 479.

If we recognize the fundamental character of this distinction, we must always keep in mind the possibility of attention directed toward the reception of the stimulus (sensorial or complete reaction), toward the execution of the movement (shortened, abbreviated, motor, or muscular reaction), or more or less evenly poised between these two attitudes (central or natural reaction).

It has also been intimated, and is apparently taken for granted by many investigators, that the length of the reaction-time is further conditioned by age, by sex, and, finally by another factor, possibly further reducible, that of constant individual differences.[1] It is this last factor which deserves our attention because it has been assumed by more than one investigator that this factor is either identical with, or at least a function of, general mental ability or general intelligence, and it has, therefore, been proposed to measure the mental ability of school-children in part by determining their reaction-time. Are such constant individual differences to be established by careful experimentation, and, if so, do they give us an indication of the general mental ability of the observers who exhibit them?

In answering these questions it seems to me that attention must be called to the methods of using the reaction-test in the psychological laboratory and in the schoolroom. The typical reaction-experiment, as conducted in the modern laboratory, is a qualitative experiment: the emphasis, I mean, is placed upon the introspective examination of the observer's consciousness rather than upon the quantitative examination of the time-values themselves. The experiment calls for the synthetic construction of the action-consciousness. The chronoscope is employed partly to supply the objective conditions of the experiment (stimulus, movement of reaction, etc.), partly to keep some tag upon the work of the observer,—in a word, to standardize the conditions under which the introspection is made. If, for example, the times recorded for a series of simple auditory reactions are 180, 170, 190, 450, and 60 sigma, it goes without saying that the observer is inaccurate in his introspective examination if he reports a uniform series of conscious processes in all five tests: something else than the simple impulsive action consciousness must have been present when 450 sigma was recorded; something must have dropped out when 60 sigma was recorded.

Again, it is customary in the conduct of the laboratory re-

[1] The very term 'personal equation' has this implication, and may possibly have suggested the use of the reaction experiment in the way we are discussing.

action-experiment (at least in the Cornell laboratory) to give heed to the so-called 'types' of Lange. An endeavor is made, first of all, to establish the 'natural' type by avoiding, in the instructions to the student, any reference to the direction of the attention either toward stimulus or movement, and especially to avoid the suggestion that the reaction is to be made as quickly as possible. Now, experience shows that, if these conditions are established, many observers will, in their first reactions, adopt the 'central' reaction: this is, then, literally, their 'natural' attitude toward the experiment.

On the other hand, rather infrequently one encounters a student who belongs very clearly either to what has been termed the 'subjective' or to the 'objective, type of observer.[1]

We have here, then, a definite indication of constant individual differences, but it is not an indication from which we can argue at all to fundamental differences in general mental ability. We find merely that these occasional observers, who are extreme in their type, fall naturally into the sensorial or into the muscular type of reactors in the absence of more definite instructions.[2]

Obviously when these observers pass to the sensorial and muscular types of reaction and are instructed specifically how to direct the attention, they find the one attitude easy and the other difficult, according as their type is objective or subjective.

But a question of further significance now arises:—can those possessed of a strong *Anlage* (whether this is regarded as a

[1] This is a distinction which has received an increasing amount of sanction in the recent experimental literature, and which has been well described by Stern in his treatise upon individual differences. It must be reducible to a fundamental difference in nervous constitution, and has, therefore, appealed to me as being possibly identical with Ribot's attractive scheme of the temperaments in which the fundamental division is into the groups,—sensitives, actives and apathetics. The sensitives may be the subjectives, the actives the objectives.

[2] To quote from Stern: "One man is accustomed to take up an active attitude to everything that occurs to him: his own action is for him the constant point of interest, and his environment is of importance only in so far as it affects this centre of reference; he therefore keeps his personality, his 'I,' in instant readiness for action. Another is wont to submit himself passively to the operation of external impressions; he regards them from the theoretical point of view, is contemplative in disposition. The former is inclined to make himself 'ready for the leap;' his fingers are intense, and the psyche is intent; he merely awaits the signal for action. The latter tends as naturally to a sensorial direction of attention: if we force him to think of himself, and to dispose himself for the coming movement before the occasion for its performance has arisen, he feels confined and confused. The former is expecting his own outbreak, the latter is awaiting the impression; the stimulus is in the one case the release, in the other the cause of movement." *Ueber Psych. d. individuellen Differenzen*, 108, as quoted by Titchener, *Exper. Psych.*, Vol. I, pt. II, 215-6.

matter of ideational type or of type of attention) be taught or coached to exhibit both the sensorial and the muscular type of reaction? I am aware that there is evidence in the negative.

Professor Baldwin has published results from which he argues that observers possessed of strong ideational types give reactions consistent with those types despite all attempts at modification, and Messrs. Hill and Watanabe, from investigations in the Cornell laboratory, conclude that "if the volitional temperament is unfavorable, practice will have no effect in determining the two types of reaction-time."[1] Despite these statements, I have personally the conviction that, at least in the vast majority of instances, a constitutional tendency toward a particular reaction-attitude can be overcome by proper coaching. In any event, we have no evidence in these laboratory investigations that there exist any constant individual differences in reaction-time that depend upon general mental ability.

With the possible exception, therefore, of a few isolated and extreme cases, we may, in the conduct of the reaction-experiment in the laboratory, bring all observers to exhibit the two chief types of attention to the reaction. And now most important is this:—when practice is attained, as indicated by a mean variation less than 10% of the average or median value, the results, judging from my experiences in the laboratory at least, are practically identical for all observers. In other words, our constant individual differences have disappeared, since they themselves depended upon degree of practice and upon variation in the direction of attention. And so we may predict with some confidence that a practiced observer, under laboratory conditions, will give the following simple reaction-times:

	SENSORIAL	MUSCULAR
pressure	210	110
sound	230	120
sight	270	180

Those few observers who cannot be taught to exhibit both types will at least conform to these figures for the type to which they belong.

Turning now to tests upon school children and students, I wish to take for illustration the work reported by Bagley, Gilbert and Wissler. It needs but little argument to show that the results of any psychophysical test are largely influenced by the conditions under which that test is administered. The

[1] For further references, see Titchener, *Exper. Psych.*, I, Part II, 225.

conditions under which the reaction-time test has been used to determine the mental ability of school children are so unlike those just described for the psychological laboratory that I venture to term them 'anthropometric' as opposed to 'laboratory' reactions. Certainly the work of the first two investigators just named belongs in the former category. For we find that emphasis is placed upon the quantitative measurement of the time to the utter neglect of the introspection of the action-consciousness, and with the avowed intent of measuring the capacity of the subject as one might measure his skull or his physical endurance. In such tests the observers are, of course, entirely unpracticed, and no specific directions are given for the control of the attention. Moreover, probably in most cases, the pupil is given to understand, either explicitly or implicitly, that the reaction is to be made *as fast as possible.* These variations in the conduct of the experiment really *make another experiment of it*, and the results cannot be expected to coincide with those obtained in the laboratory. Hence the question must always be asked: how far are such tests reaction-times in the strict sense of the term, and, further, just what is actually being measured.

Bagley[1] attempted to relate mental and motor ability by the use of both experimental and non-experimental values. The experimental data for mental ability were the results of reaction-tests made with the aid of Jastrow's card-sorting apparatus.[2] These reactions are not, of course, simple reactions, but the method lends itself, nevertheless, to discussion in this place.

It will be noted that Bagley assumed definitely at the outset that these reaction-times would be expected to indicate mental ability. He says (p. 194): "the experimental sources of the data on mental efficiency consisted of various types of reaction-times as representing quantitatively the mental ability of the subject, mental excellence being represented by the alertness of the mind in reacting appropriately to given stimuli."

Glancing at his results, we find, however, that, by the method of correlation by groups, there is a discrepancy between reaction-times and class-standing, *i. e.*, the relation is not positively inverse, but indifferent. There follows, moreover, this significant clause: "the children whose reaction-times are nearest the norm have the best class-standing, while those who are particularly alert and those who are particularly slow in

[1]On the Correlation of Mental and Motor Ability in School Children, this *Journal*, XII, 193.

[2]It may be pointed out parenthetically that card sorting is open to numerous objections. In particular, the variable factor of the 'handling-time' cannot be adequately eliminated by mere subtraction, as I have convinced myself by personal trial.

reaction, are alike deficient in mental efficiency as represented by the class-standing." Now, would it not be plausible to modify this inference in this way:—the most intelligent children, as indicated by class-standing, are most able to follow instructions, and therefore to approximate the norm, while the less attentive children are erratic and prone to yield either premature or delayed reactions? This, it seems to me, is the true solution of Bagley's results.

Gilbert[1] tested school children of various ages, using a visual discrimination reaction (react to blue, not to red) and also a simple visual reaction. The results of the two tests were closely similar and may be considered together.

In his recent article upon the objective measurement of "General Intelligence,"[2] C. Spearman criticises Gilbert for not working out the correlation between reaction-time and mental ability by more exact formulæ, and explains the low figure which can thus be actually found as the result of the unfavorable conditions under which the procedure was made, the heterogeneity of the reagents, etc., with the final conclusion that the true correlations must be very much higher than those actually observed.

It seems to the writer that such criticism fails to reach the heart of the matter. How would Gilbert's results stand in the face of the rigorous standardization of conditions which we have found necessary in the laboratory? Is there any value in computed correlations from data which are themselves so loosely gathered?

For instance, it will be recalled that Gilbert argues at length for the use of the median-value versus the arithmetical average. He says that in a series such as 240, 230, 220, 250, 480, 430 sigma, etc., the median-value method is more satisfactory because the large variants, like 480 and 430, do not so much influence the final value, while their force is directed, as it should be, toward increasing the size of the mean-variation. But in the light of all we know of the conduct of the experiment in the laboratory, such variants as 480 and 430 ought not to be admitted to calculation under *any* method, for they are clearly not *bona-fide* measurements of the reaction-consciousness: they do not stand for, or measure, the same psychophysical process as the other values cited. Why, then, should they be classed with them for further treatment? In the laboratory they would be discarded,—not by the experimenter, but by the observer's introspective control. Inspection of Gilbert's tables shows that the median value is, well-nigh uniformly,

[1] *Studies Yale Psych. Lab.*, II, Nov., 1894, 40.
[2] This *Journal*, XV, 201 ff., esp. 280-1.

from 15 to 20 sigma shorter than the arithmetical average: this indicates the presence of a few large values like those mentioned in nearly every group of tests.

Again, suppose we compare the actual times with those given in the laboratory. The values for the older children range, in some groups, from 170 to 180 sigma (arithmetical) and from 155 to 172 sigma (median value). These values, then, are *less* than the *muscular* reaction-times of practiced laboratory observers. But can we stop here? Can we even suppose that Gilbert's school children were mainly giving muscular reactions? Despite the possible suggestion of reacting as quickly as possible, I feel sure that this was not the case, and for three reasons:

(1) The rapid, semi-automatic reaction is nearly always the result of practice and is not apt to appear in a short first series.

(2) It is well known that visual reactions, more than those of any other modality, favor the sensorial type in an unpracticed observer, on account of the necessity of sensory accommodation for the reception of the stimulus.

(3) Gilbert took express pains to foster the sensorial attitude in order to guard against premature reactions. This he did by placing the discrimination reaction before the simple reaction in the order of tests, in order thereby to engender a habit of watching the stimulus.

We may, therefore, look with some suspicion upon such curiously small average times as those recorded by Gilbert.

And, again, if we add to this that Gilbert's mean variations are far above the 10% limit allowed by the laboratory, we are assured that such values are practically useless as measurements of reaction-times. They illustrate very well what I have in mind as the divergence of the anthropometric from the laboratory standards of conducting psychological tests.

But, even assuming that these values represented trustworthy reactions, what evidence have we that they are correlated with mental ability? When we are told that 'bright' children react more quickly than 'dull,' though less difference can be made out in the case of the 'average' group, we find the quantitative support is afforded by the figures 207, 213 and 224 for the three groups. And how much within the limits of certainty are we when the statement that "the brighter the child the more quickly he is able to react with discrimination and choice" is supported by the values 401, 402 and 420 sigma for bright, average and dull children respectively?

From a gradual increase in the rapidity of the reaction with age, even if established, we cannot argue that the increase is attributable to increased general mental ability. Mental growth and mental maturity must never be confounded with mental ability.

The work of Wissler[1] presents an exhaustive treatment of the results of various physical and mental tests of students of Columbia University as compared with one another and with the class standings. The conditions under which reactions were taken in this series are evidently midway between the best laboratory conditions and the looser conditions of Bagley and Gilbert. The reagents are very much of an age, are presumably all of intellectual ability above the average, while the external objective conditions were those used in the psychological laboratory. On the other hand there is no record of introspective control. The tests were few in number (three to five from a set of five), and evidently somewhat hurriedly taken, as Spearman has surmised.[2] Wissler, however, shows clearly that reaction-times are useless as indications of mental ability or even, what is more surprising, of general alertness or readiness of thought or action. The marking of A's, the naming of colors, the execution of rapid movements, and the recording of simple associations (all tests which put a premium upon rapidity and general psychophysical facility) fail to show any significant correlations with one another,—a conclusion, as the author remarks, "out of harmony with all general belief." And the relation of reaction-times with class-standings is merely that of chance.

If our arguments have not gone astray, we seem bound, in the light of our discussion, to conclude that the reaction-test is quite without significance as a measure of mental ability (save in so far as a small mean variation might indicate a certain steadiness in the control and direction of attention). We see that the reaction-time of any observer is determined by a large number of more or less independent factors, and that, when these factors have either been eliminated or controlled, as they are in laboratory procedure, we have left no residuum of individual variation that can be turned to account in estimating the observer's general intelligence or mental ability.

May it not be, or rather, is it not certain, that many other psychophysical tests which have been employed in a rough and ready way, often by inexpert experimenters, especially in the tests conducted upon school children, are quite as likely to be vitiated by factors similar to those we have considered?

At the risk of a digression, I may cite, as an instance, the test of pitch-discrimination which Spearman[3] has just advocated as a universal and ready means for the estimation of "General Intelligence." After conducting tests in pitch-discrimination in the psychological laboratory under all sorts of con-

[1] The Correlation of Mental and Physical Tests, *Mon. Suppl. Psych. Rev.*, June, 1901.

[2] *Loc. cit.*, 283.

[3] *Loc. cit.*

ditions for more than three years, I am positive that the use of a monochord to measure pitch-discrimination, as was done by Spearman, is far too rough a method to use for this purpose. What, for example, shall we infer in regard to the general intelligence of the young lady reported by me in a previous study[1] who was totally unable to judge a half-tone when struck with the slightest unevenness of accent upon the piano, but who could, after a little practice, give a value of S = 0.52 vibs. when the Stern tone-variator was used, and the influence of disturbing secondary factors thereby eliminated? I cannot forbear quoting my previous summary of these tests of pitch-discrimination: "A typical unmusical observer, when placed under proper conditions, may discriminate pitch differences of less than three vibrations correctly in 75% of the tests, but if the stimuli are of relatively low pitch, if they are given without any preliminary 'warming-up,' if the time-interval between them exceeds four seconds, if they are given too briefly or in too quick succession, if they are of unequal intensity, or if they are presented simultaneously with one or more other similar stimuli, then discrimination becomes either difficult or quite impossible, and it may then remain impossible even when D is represented not by a few vibrations, but by musical intervals of one or two octaves and more."

We have here, evidently, simply another illustration of the proposition which I have attempted to establish in regard to the reaction-experiment, viz:—the results of a psychophysical test are determined very largely by the conditions under which the test is conducted.

SUMMARY.

(1) The reaction experiment, as conducted in the psychological laboratory, is primarily a qualitative experiment for the synthesis of the action-consciousness. When due regard is paid to the objective conditions, to practice, and to the direction of attention, most observers can readily be trained to adopt any of the three types of attention. And, for a given direction of the attention, the quantitative results are uniform for all observers.

(2) The reaction experiment of the 'anthropometric' type, as conducted upon school children, is a rough and ready quantitative experiment, which has yielded results of very questionable value as adjudged by the laboratory standards. The constant individual differences between groups of children, which have been assumed in some studies to indicate differing grades of mental ability or general intelligence, are largely the products of faulty experimental conditions, and cannot, in any event, be

[1] This *Journal*, XIV, 297-304.

referred, either in theory or practice, to constant individual differences in mental ability.

(3) This examination of the reaction-experiment suggests that many other experimental studies of school children have been vitiated by the neglect of those cautions in procedure and interpretation that laboratory practice has taught. As an illustration, the pitch-discrimination of an unmusical observer is shown to be subject to wide variation when the test-conditions are modified but slightly.

(4) The outcome of the reaction-time test (and, indeed, of any psychophysical test) upon school children will, furthermore, depend not only upon the objective conditions of the test, upon the nature of the instructions given, etc., but also to an appreciable extent upon the ability of each child to understand and carry out these instructions. When, therefore, a test is affected in this way, any assumed correlation between the quantitative results and the general intelligence of the group of children tested is, in reality, but a correlation of general intelligence with itself.